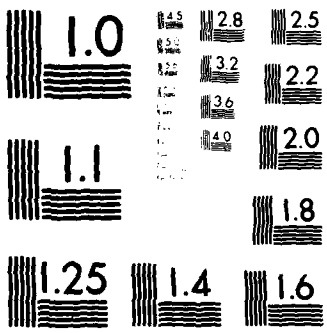
AD-A170 239 ON NONGAUSSIAN SIGNAL DETECTION AND CHANNEL CAPACITY (U) NORTH CAROLINA UNIV AT CHAPEL HILL DEPT OF STATISTICS C R BAKER AUG 85 N00014-81-K-0373

1/1

UNCLASSIFIED F/G 9/4 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

(12)

AD-A170 239

On NonGaussian Signal Detection and Channel Capacity

Charles R. Baker

Department of Statistics
University of North Carolina
Chapel Hill, N.C. 27514

LISS 14

August, 1985

Research supported by ONR Contracts N00014-75-C-0491 and N00014-81-K-0373

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release: Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Department of Statistics — 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): University of North Carolina, Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research — 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-84-C-0212; N00014-81-K-0373 & N00014-75-C-0491
8c. ADDRESS (City, State and ZIP Code): Statistics & Probability Program, Arlington, VA 22217	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
NR	042	269	SRO 105

11. TITLE (Include Security Classification): On NonGaussian Signal Detection & Channel Capacity

12. PERSONAL AUTHOR(S): Charles R. Baker

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
TECHNICAL	FROM ____ TO ____	Aug., 1985	18

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	signal detection; nonGaussian stochastic processes; Shannon theory; channel capacity; likelihood ratio

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A summary of recent research results is given. Three problems are discussed: detection of nonGaussian signals in Gaussian noise, signal detection in nonGaussian spherically-invariant noise, and information capacity of communication channels when the constraint on the transmitted signal is mismatched to the Gaussian channel noise.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL: C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code): (919) 962-2189 — 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

ABSTRACT

A summary of recent research results is given. Three problems are discussed: detection of nonGaussian signals in Gaussian noise, signal detection in nonGaussian spherically-invariant noise, and information capacity of communication channels when the constraint on the transmitted signal is mismatched to the Gaussian channel noise.

A-1

ON NONGAUSSIAN SIGNAL DETECTION AND CHANNEL CAPACITY

Charles R. Baker*
Department of Statistics
University of North Carolina
Chapel Hill, NC 27514

INTRODUCTION

This paper contains a discussion of some recent research in signal detection and communications. No proofs are included; the emphasis is on motivation and results. Some precise definitions and results are contained in the Appendix. The paper is couched in terms of problems in underwater acoustics; as will be seen, the models and results are of general applicability.

SIGNAL DETECTION

Two problems will be discussed. The first, for which the most complete results were obtained, was that of detecting nonGaussian signals in Gaussian noise. The second, motivated by examination of noise properties for actual sonar data, was that of detection in a class of nonGaussian processes, which can be regarded as mixtures of Gaussian processes.

Both of these problems are important in sonar applications. The detection of nonGaussian signals in Gaussian noise can be regarded as the canonical problem for active detection in reverberation-limited noise, especially volume reverberation. The noise in such situations can frequently be regarded as arising from reflections by many small scatterers, which can be reasonably assumed to have statistically-independent behavior. The central limit theorem then gives a Gaussian process. The signal process, however, will frequently be dominated by reflections from a few large scatterers, such as

* Research supported by ONR Contracts N00014-75-C-0491, N00014-81-K-0373, and N00014-84-C-0212.

the sonar dome. These scatterers each give rise to a nonGaussian random process, which are summed at the receiver to give a nonGaussian process.

Other applications may also involve detection of nonGaussian signals in Gaussian noise. For example, an emerging passive sonar detection problem is that of detecting very quiet submarines, emanating primarily broadband signals. These signals may prove to be nonGaussian and one will frequently be faced with detecting them in a Gaussian noise background.

The importance of the problem of detecting signals in nonGaussian noise of a "spherically-invariant" (Gaussian mixture) type has become apparent by examining the results of data analysis on acoustic recordings obtained from both under-ice and shallow-water environments. These noise recordings have exhibited data whose univariate distribution properties appear similar to those of Gaussian random variables (symmetric, unimodal, smooth). However, when compared with zero-mean Gaussian random variables having the same variance, the data often exhibits heavy tails and/or high kurtosis. These features, at least in the univariate case, are very appropriate to a Gaussian mixture model for the noise. If the multivariate data has a Gaussian mixture distribution, then the modeling problem (in the context of modeling nonGaussian processes) is greatly simplified, as will be discussed below.

DETECTION OF NONGAUSSIAN SIGNALS IN GAUSSIAN NOISE

Our objective here was to give a complete solution of the detection problem. The actual data processes appear as functions of continuous time. The desirable results then include the following:

(1) Characterization of signal-plus-noise processes for which the detection problem is well defined. By this, we mean a mathematical model which does not promise perfect (singular) detection. Such singular models are not considered to be realistic.

(2) For well-defined problems, determination of the likelihood ratio for the continuous-time problem.

(3) Approximation of the continuous-time likelihood ratio by a discrete-time form, preferably in recursive or near-recursive form.

(4) Specification of procedures for estimating parameters appearing in the approximation to the likelihood ratio.

(5) Performance evaluation of the approximation to the likelihood ratio.

Of course, there are other desirable results, such as development of robust approximations to the likelihood ratio, and approximations which do not require a full description of data parameters. However, the results listed in (1)-(5) above are already very ambitious, and obtaining them would be a significant step in any complete solution to the problem.

Considerable work has been done on detection of Gaussian signals in Gaussian noise; see for example some of the references given in [5]. However, previous work on detection of nonGaussian signals in Gaussian noise has been subject to one or more of the following limitations:

(a) The noise is assumed to be the Wiener process; see, e.g., [14]. The paths of the Wiener process are far too irregular to reasonably model sonar noise. Other Wiener process properties, such as independent increments, Markov, etc., are not typically satisfied. Moreover, determination of likelihood ratio parameters is left as an open problem.

(b) Detection is based on second-moment criteria, such as the deflection criterion [1], [2].

(c) Signal and noise are taken to be independent [3], and expressions for the likelihood ratios are not obtained.

We have obtained a reasonably complete solution to problems (1)-(3) above. The solutions to (4) and (5) are presently being computationally investigated. Here we give a rough summary of the results. More precise statements require a substantial mathematical machinery; reference is made to [8] for the complete and final statements; partial

results are contained in the Appendix.

The data is assumed to be observed over a finite interval, which we take as [0,1]. The noise is Gaussian, mean-square continuous, zero-mean, and is assumed to vanish (almost surely) at t=0.

To solve problem (1) mentioned above, one would wish to consider a general nonGaussian process (Y_t), and determine necessary and sufficient conditions for the detection problem to be well-defined (non-singular). Such conditions are given in Theorem 1 and Theorem 2 of the Appendix. Roughly, they require that the process (Y_t) have a signal-plus-noise representation $Y_t = S_t + N_t$, where the sample paths of (S_t) belong almost surely to the reproducing kernel Hilbert space of (N_t). This condition means that the covariance function (resp., sample paths) of the signal process must be much smoother than the covariance (resp., sample paths) of the noise. The process (S_t) must also satisfy certain measurability conditions with respect to (Y_t) and (N_t). See the Appendix for the precise statements. We remark that the necessary conditions and the sufficient conditions are not identical, although they are very close.

Problem (2) mentioned above is that of determining the continuous-time likelihood ratio. A general solution has been obtained, and is given in [8]. Actually, two solutions are given there. One views the observations as being simply real-valued functions; the other treats them as being elements of $L_2[0,1]$. The latter is summarized in the Appendix. Here we shall give the finite-sample discrete-time approximation to the likelihood ratio on $L_2[0,1]$.

First, the noise is a Gaussian vector $\underline{N}$ having covariance matrix $\underline{R}$. We can represent $\underline{R}$ by $\underline{R} = \tau^2 \underline{F}\underline{F}^*$ where $\underline{F}$ is a lower-triangular matrix, and τ is the sampling interval. We can thus consider the noise process to be a sampled version of $N_t = \int_0^t F(t,s)dW_s$, where (W_s) is the standard Wiener process. According to the results

of [8], the signal-plus-noise process will be of the form $S_t + N_t = \int_0^t F(t,s)dZ_s$, where (Z_t) is here taken to be a diffusion with memoryless drift function σ: $Z_t = \int_0^t \sigma_s(Z_s)ds + W_t$. The resulting discrete-time approximation to the log-likelihood ratio is then

$$\Lambda^{n+1}(\underline{X}^{n+1}) = \Lambda^n(\underline{X}^n) + \tau^{1/2}\sigma_n[\tau^{1/2}(\underline{L}\underline{E}_n^{-1}\underline{X}^n)_n](\underline{E}_{n+1}^{-1}\ \underline{X}^{n+1})_{n+1}$$
$$- \frac{1}{2}\tau\sigma_n^2[\tau^{1/2}(\underline{L}\underline{E}_n^{-1}\underline{X}^n)_n], \quad n \geq 1;$$
$$\Lambda^1(\underline{X}^1) = 0$$

where $\underline{X}^n$ denotes the observation vector obtained from the first n samples;

$\tau^2\, \underline{E}_n\, \underline{E}_n'$ is the noise covariance matrix for the first n sample times;

$\underline{L}$ is the summation matrix: $(\underline{L}\underline{X}^n)_i = \sum_{j=1}^{i} X_j^n$.

This formulation of the log-likelihood ratio is partially recursive. Note that

$$(\underline{L}\underline{E}_n^{-1}\underline{X}^n)_n = (\underline{L}\underline{E}_{n-1}^{-1}\underline{X}^{n-1})_{n-1} + (\underline{E}_n^{-1}\underline{X}^n)_n,$$

and that the operation $(\underline{E}_n^{-1}\underline{X}^n)_n$ is just a cross-correlation of the data vector with the n^{th} row of $\underline{E}_n^{-1}$.

There are three basic considerations in evaluating the usefulness of the above log-likelihood ratio. One is the validity of the approximation assumption; a second is the development of procedures for estimating the parameters of the likelihood ratio; finally, one is interested in whether or not the discrete-time approximation is in fact a likelihood ratio when our assumptions are satisfied. We discuss these three points below.

(i) Any Gaussian vector can be obtained by passing white Gaussian noise through an appropriate lower-triangular matrix. Thus, the noise model is reasonable for the discrete-time problem, and one can justify the use of multiplicity M=1 from this and from other mathematical considerations. The fact that (Z_t) is a process of diffusion type then follows from well-known results [14]; to assume further that it is of diffusion type with respect to (W_t), one reasons that the difficult detection problems are of most interest; such problems are those in which the N and $S+N$ processes have very similar properties. Since (N_t) is modeled as a time-varying linear operation on the diffusion (W_t), it seems reasonable to model (Y_t) as that

same time-varying linear operation on a process that is of diffusion type with respect to (W_t). More detailed physical interpretations of the assumptions can be given for applications in sonar. However, a basic reason for making the assumptions is that they permit one to implement an approximation to the likelihood ratio without detailed knowledge of the data probability distributions. The validity of these assumptions and the effectiveness of the finite-sample discrete-time approximations can be judged in each application by the performance of the detection algorithm.

(ii) The implementation of the sequence of test statistics (Λ^k) given above, for $k \leq n$, requires knowledge of only two parameters: the lower-triangular matrix $\underline{F}_n$ such that $\tau^2 \underline{F}_n \underline{F}_n^*$ is the $n \times n$ noise covariance matrix, and the drift function σ. Typically (in sonar) these quantities will need to be estimated from experimental data. We give a procedure for doing this, supposing that one has an ensemble of independent sample vectors from the noise process, of sufficient size to give a good estimate of the covariance matrix, and that one or more sample vectors from the signal-plus-noise process is available.

First, the noise vector is written as $\underline{N} = \underline{F}\, \underline{\Delta W}$, where ΔW is the vector with j^{th} component $(W[j\tau]-W[(j-1)\tau])$ for $j \geq 2$, and first component $W(\tau)$. If $i\tau = t_i$ and the ij element of $\underline{F}$ is $F(i,j)$ for all ij, then $N_i \cong N(i\tau) = N(t_i)$ for large i and small τ. The representation for $\underline{N}$ gives noise covariance matrix $\underline{R}_N = \tau^2 \underline{F}\underline{F}^*$. Consistent with this representation and the results of [8], the $S+N$ vector is written as $\underline{Y} = \underline{F}\underline{\Delta Z}$, where $\underline{\Delta Z}$ is the vector with j^{th} component $(Z[j\tau]-Z[(j-1)\tau])$ for $j \geq 2$, and first component $Z(\tau)$. Thus, given an ensemble of sample noise vectors, one treats the resulting estimate of the noise covariance matrix as $\underline{R}_N$, and obtains the factorization $\underline{R}_N = \tau^2 \underline{F}\underline{F}^*$. Then, given a $\underline{Y}$ sample vector, $\underline{Z}$ is estimated by $\underline{\Delta Z} = \underline{F}^{-1}\underline{Y}$ and $(\underline{\Delta Z})_1 = Z(\tau)$. Given $\underline{Z}$, our assumptions yield

$$Z_i = Z(i\tau) = \int_0^{i\tau} \sigma_s(Z_s)ds + W(i\tau)$$

Various methods can then be used to estimate the unknown function σ. A general maximum-likelihood estimate is given in [11], which is now being computationally investigated. The generality of this procedure leads to computational difficulties, so that it may be necessary to assume a specific form for σ, such as a low-order polynomial with unknown coefficients.

(iii) Under our assumptions, Λ^n can be considered a "good" approximation to a likelihood ratio test statistic if τ is "small" and n is "large ". These are not exact statements; at present, we have no bounds on performance. Any such bounds would involve $\underline{R}_N$, τ, σ, and n. However, if $\underline{Y}$ is Gaussian, a precise statement can be made. Suppose that $\underline{N} = \underline{F}\underline{\Delta W}$ and $\underline{Y} = \underline{F}\underline{\Delta Z}$ as above, and that (a_k) and (b_k) are two sequences of real numbers such that

$$Z^k = \sum_{j=1}^{k-1} [a_j Z_j + b_j] + W_k \qquad 2 \leq k \leq n,$$

$$Z_1 = W_1.$$

In this case, it can be shown that exp (Λ^n) is a monotone function of $dP_{\underline{Y}}/dP_{\underline{N}}$, thus a likelihood ratio test statistic. We conjecture that this also holds when $\underline{Y}$ is

not Gaussian:

$$N = \underline{F} \underline{\Delta W}, \underline{Y} = \underline{F} \underline{\Delta Z}, \text{ and}$$

$$Z_k = \sum_{j=1}^{k-1} \sigma_j(Z_j) + W_k \qquad 1 \leq k \leq n,$$

with σ non-affine.

The above results give a solution to two problems of much interest in the theory of stochastic processes: determining conditions for a discrimination problem to be non-singular, and determining the likelihood ratio, when one of the two processes is Gaussian. The scope of these problems can be appreciated by reviewing some of the references cited in [5]. The above approximation to the log-likelihood ratio gives some hope of obtaining useful new detection algorithms for some important sonar detection problems. The eventual utility, however, will be apparent only after a great deal of further work is done, especially computational work involving experimental data.

DETECTION IN NONGAUSSIAN NOISE

Examination of data properties by several investigators has indicated that sonar data may be spherically-invariant (a Gaussian mixture) in several important applications. One such application is in under-ice operations. Analysis of such data has shown that the univariate data typically has high kurtosis and heavy tails as compared to Gaussian data of the same variance [10].

Another environmental situation which apparently gives rise to univariate spherically-invariant noise is that of near-shore operations in warm climes (e.g., the Gulf of Mexico). The nonGaussian noise in this case is attributed to snapping shrimp and results in very high kurtosis as compared to Gaussian data [15].

These observed data properties motivated us to consider the problem of detection in spherically-invariant noise. Such a noise process (N_t) can be represented as $N_t = AG_t$, where (G_t) is a zero-mean Gaussian process with covariance function R, and

A is a random variable independent of (G_t). Such processes are also said to be Gaussian-mixture processes.

Of course, the property of being univariate spherically-invariant does not imply that a process will be spherically-invariant in the multivariate case, as consideration of the case $A = 1$ will show. However, if the above representation of the noise is reasonable, then the problem of characterizing the probability distributions for nonGaussian noise is reduced to that of determining the covariance function of (G_t) and the probability distribution function of the random variable A. Without loss of generality, one can assume that the second moment $EA^2 = 1$, so that the covariance of (G_t) is the same as that of (N_t). As this function can be estimated, the major problem is that of determining the distribution function of A. We are presently carrying out computational work on this problem, using maximum-likelihood estimation.

The significance of this model, if accurate, is that it would permit one to describe all the joint distributions of the data through knowledge of the covariance (as in the Gaussian case) and of the distribution function for a single random variable. It can thus be viewed as a first step away from the Gaussian noise hypothesis which does not require that one take independent samples.

We are interested here in obtaining the same results (1)-(5) discussed above for the Gaussian noise case. So far, partial results have been obtained for (1) and (2). We have found [9] that the sufficient conditions for a well-defined (non-singular) detection problem are the same as those obtained for detection in Gaussian noise (which are very close to being necessary). An expression for the continuous-time likelihood ratio has also been found. The remaining problems in obtaining (3)-(5) above have yet to be seriously investigated.

MUTUAL INFORMATION AND CHANNEL CAPACITY

Work to be discussed here was again for two types of noise processes: one where the channel noise is Gaussian, the other where it is spherically-invariant.

The capacity of a communication channel is here taken to be its information capacity:

$$C = \sup_{Q} I(m, Y)$$

where m is the message process, $A(m) = s$ is the transmitted signal, N is the channel noise, $Y = A(m) + N$ is the received process, and $I(u, v)$ is the mutual information between stochastic processes u and v (as defined in [4]). The constraint class Q contains all admissible message processes m and coding functions A. It is usually chosen from considerations involving average power, so typically involves a relation between the signal process and the noise covariance. In the case of stationary signal and noise processes, with spectral density functions Φ_s and Φ_N, an appropriate constraint is

$$\int_{-\infty}^{\infty} \frac{\Phi_s}{\Phi_N}(\lambda) d\lambda \leq P.$$

This can be related to the reproducing kernel Hilbert space of the noise process, and a related general constraint is $E\|A(m)\|_N^2 \leq P$, where $\|u\|_N$ is the reproducing kernel Hilbert space (for N) norm of the function $u(t)$.

If one considers this in physical terms for the frequency domain, such a constraint places a limitation on the expected value of the integrated ratio of signal energy to noise energy. In non-white noise, this is obviously more realistic than a limitation on total signal energy alone.

MISMATCHED CHANNELS

With the type of constraint discussed above, a complete solution to the channel capacity problem for Gaussian channels without feedback is given in [4]. However, this approach will not be valid when the covariance of the channel noise (N_t) is unknown.

This can occur from natural causes, as with insufficient knowledge of the environment. It can also occur because of jamming in the channel. In the latter case, it is well-known that if the channel noise has a given covariance, then channel capacity is minimized when the noise is Gaussian. Thus, a jammer seeking to minimize capacity of a channel with ambient Gaussian noise would choose to add Gaussian noise, and the channel capacity would then be determined by the relation between the actual channel noise (including the jammer's contribution) and the noise covariance assumed by the user of the channel. Of course, less obvious questions also arise. Channel capacity is only the starting point in analyzing such situations.

These considerations have motivated us to introduce the notion of "mismatched" channels, wherein the constraint on transmitted signals is taken with respect to a covariance which is different from that of the channel noise.

An analysis of this problem is contained in [6] for a large class of Gaussian noise processes. Additional results are forthcoming [7]. Striking differences appear between the results for the mismatched channel and those for the matched channel (when the channel noise is also the constraint noise). For example, in the matched continuous-time channel with the above generalized power constraint ($E\|A(m)\|_N^2 \leq P$), the capacity of the channel is equal to $P/2$ and cannot be actually attained. In the mismatched channel, the capacity can be either greater or smaller than the capacity for the matched channel and it can be attained in some situations. The value of the capacity depends on the relation between the two covariances. We give one result from [6].

Let the constraint covariance operator in $L_2[0,T]$ be denoted by R_W (consider W as the noise assumed by the channel user) and let R_N be the covariance operator for the channel noise process N. Suppose that $R_N = R_W^{1/2}(I+S)R_W^{1/2}$ where I is the identity in $L_2[0,1]$ and S is a compact operator. This relation will be satisfied, for example, when

W and N are two Gaussian processes for which the discrimination (detection) problem is well-defined. Let

$$C_W(P) = \sup_Q I[A(m), Y]$$

when Q contains all coding operations A and stochastic processes m (including nonGaussian processes) on [0,T] such that $E\|A(m)\|_W^2 \leq P$. Finally, let $\{\lambda_n, n \geq 1\}$ denote the strictly negative eigenvalues of the operator S defined above. Of course, this set may be empty, as when N can be written as $N = W+V$, with V independent of W. Let $\{e_n, n \geq 1\}$ be associated o.n. - eigenvectors. Then [6]:

(a) If $\{\lambda_n, n \geq 1\}$ is not empty and $\sum_n |\lambda_n| \leq P$, then $C_W(P) = \frac{1}{2} \sum_n \log[(1+\lambda_n)^{-1}] + \frac{1}{2}[P + \sum_m \lambda_m]$.

(b) If $\{\lambda_n, n \geq 1\}$ is not empty, and $\sum_n |\lambda_n| > P$, then there exists a largest integer K such that $\sum_1^K \lambda_i + P \geq K\lambda_K$, and

$$C_W(P) = \frac{1}{2}\sum_{n=1}^{K} \log\left[\frac{\sum_1^K \lambda_i + P + K}{K(1+\lambda_n)}\right]$$

(c) If $\{\lambda_n, N \geq 1\}$ is empty, $C_W(P) = P/2$.

(d) In (a) and (b), the capacity is strictly greater than when $R_N = R_W$; in (c) these capacities are equal.

(e) In (a), the capacity can be attained if and only if $\sum_n |\lambda_n| = P$. It is then attained by a Gaussian signal with covariance operator $R = \sum_{i=1}^K \beta_i u_i \otimes u_i$, where $u_n = Ue_n$, U unitary, and

$\beta_n = -\lambda_n(1+\lambda_n)^{-1}$ for $n \geq 1$. In (b), the capacity can be attained by a Gaussian signal process with covariance operator as above, with $u_n = Ue_n$ and

$\beta_n = (1+\lambda_n)^{-1}[\sum_1^K \lambda_i + P + K]/K$ for $n \leq K$;

$\beta_n = 0$ for $n > K$. In (c), the capacity cannot be attained.

A more general model is considered in [7]. That model is for the case where S is not necessarily compact, but has a pure point spectrum. The capacity for the mismatched Gaussian channel can then be either smaller or larger than that of the matched channel, depending on the spectral properties of the operator S.

CAPACITY OF SPHERICALLY-INVARIANT CHANNELS

The apparent usefulness of a spherically-invariant process to model noise in under-ice and shallow-water applications has motivated us to examine the channel capacity problem for communicating in such noise. This work has been aided by the work on signal detection described above; in fact, the likelihood ratio plays a key role in channel capacity problems.

We have examined the problem for the matched channel, where the constraint on transmitted signal is $E\|A(m)\|_N^2 \leq P$, with N the channel noise. As shown in [4] and [13], the capacity for the matched Gaussian channel with this constraint is $P/2$, with or without feedback. For the spherically-invariant channel with noise model $N_t = AG_t$, A a random variable independent of the Gaussian process (G_t),$EA^2 = 1$, we have found the capacity to be equal to $\frac{P}{2}E(A^{-2})$. $E(A^{-2})$ will typically be quite large for some underwater acoustics applications. Thus, this result holds forth the possibility that one may be able to communicate at much higher rates than for the Gaussian channel with the same covariance.

REFERENCES

(1) C.R. Baker, "Optimum quadratic detection of a random vector in Gaussian noise", *IEEE Trans. on Communications Technology, 14,* 802-805 (1966).

(2) C.R. Baker, "On the deflection of a quadratic-linear test statistic", *IEEE Trans. on Information Theory, 15,* 16-21 (1969).

(3) C.R. Baker, "On equivalence of probability measures", *Annals of Probability, 1,* 690-698 (1973).

(4) C.R. Baker, "Capacity of the Gaussian channel without feedback", *Information and Control, 37,* 70-89 (1978).

(5) C.R. Baker, "Absolute continuity of measures on infinite-dimensional linear spaces", *Encyclopedia of Statistical Sciences, Vol. 1,* John Wiley & Sons, 3-11 (1982).

(6) C.R. Baker, "Channel models and their capacity", *Essays in Statistics: Contributions in Honour of Norman L. Johnson,* 1-16, P.K. Sen, ed. (North Holland, 1983).

(7) C.R. Baker, "Capacity of mismatched Gaussian channels", *IEEE Trans. on Information Theory* (to appear).

(8) C.R. Baker and A.F. Gualtierotti, "Discrimination with respect to a Gaussian process", *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete.*

(9) C.R. Baker and A.F. Gualtierotti, "Signal detection and channel capacity for spherically-invariant processes, *Proceedings 23rd IEEE Conference on Decision and Control,* 1444-1446 (1984).

(10) R.F. Dwyer, "A technique for improving detection and estimation of signals contaminated by under-ice noise", in *Statistical Signal Processing,* E.J. Wegman and J.G. Smith, eds., 153-166, Marcel Dekker (1984).

(11) S. Geman, "An application of the method of sieves: functional estimator of the drift of a diffusion", *Colloq. Math. Soc. Janos Bolyai,* 32, (Nonparametric Statistical Inference), Budapest (1980).

(12) T. Hida, "Canonical representations of Gaussian processes and their applications", *Mem. Coll. Science, Kyoto University,* 33A, 109-155 (1960).

(13) M. Hitsuda and S. Ihara, "Gaussian channels and the optimal coding", *J. Multivariate Analysis, 5,* 106-118 (1975).

(14) R.S. Liptser and A.N. Shiryayev, *Statistics of Random Processes I: General Theory,* Springer-Verlag (1977).

(15) G.R. Wilson and D.R. Powell, "Experimental and model density estimates of underwater acoustic returns", in *Statistical Signal Processing,* E.J. Wegman and J.G. Smith, eds., 223-240, Marcel Dekker (1984).

APPENDIX

Absolute Continuity and Likelihood Ratio

Definitions and Notation

All stochastic processes are defined on the probability space (Ω,β,P), with parameter set $[0,1]$. R^K is the Borel σ-field for R^K, $K<\infty$, $C_0[0,1] \equiv C_0$ is the set of all real-valued continuous functions on $[0,1]$ that vanish at zero. C is the Borel σ-field on C_0 defined by the sup norm. C^K is the Borel σ-field of C_0^K under the product topology; C_0^K can be identified with the set of all K-component real-valued vector functions having each component in C_0.

Suppose that (V_t) is a vector stochastic process such that $V(\omega,\cdot) \in C_0^K$ a.e. $dP(\omega)$. V will denote the corresponding path map from Ω into C_0^K, and P_V the induced measure on $C^K : P_V = P \circ V^{-1}$.

(N_t) will denote the noise; it is m.s.-continuous, Gaussian, zero-mean, and vanishes at t=0 w.p. 1. (N_t) is thus purely deterministic, so has a proper canonical Cramer-Hida representation [12]:

$$N_t = \sum_{i=1}^{M} \int_0^t F_i(t,s)dB_i(s) \qquad (1)$$

where $M \leq \infty$ is the multiplicity of (N_t), each F_i is a deterministic Volterra kernel, and the B_i's are mutually orthogonal stochastic processes with orthogonal increments. (N_t) is Gaussian; the B_i's are thus mutually independent Gaussian processes with independent increments and continuous variances. Each B_i is thereby path-continuous. Since the representation (1) is proper canonical, and (N_t) and the family of B_i's are Gaussian, the σ-field generated by $\{N_u, u \leq s\}$ is the same as the σ-field generated by $\{B_i(u), u \leq s, i \leq M\}$, for all s in $[0,1]$. β_i will denote the Borel measure on $[0,1]$

defined by the continuous non-decreasing variances EB_s^2, $0 \leq s \leq 1$.

We assume that the multiplicity M of (N_t) is finite. This restriction is due to the absence of some needed results in infinite-dimensional stochastic calculus. However, based on a partial investigation, we believe that the results on absolute continuity and likelihood ratio presented here remain valid for $M = \infty$.

Suppose that (V_t) is any stochastic process; $\underline{\sigma}(V)$ is the P-completed filtration generated by (V_t), and $\underline{\sigma}(V) \vee \underline{\sigma}(N)$ is the smallest filtration containing both $\underline{\sigma}(V)$ and $\underline{\sigma}(N)$. We recall that a process (X_t) is $\underline{\sigma}(V)$-predictable if $G:(t,\omega) \rightarrow X_t(\omega)$ is measurable with respect to the predictable σ-field $P(V)$ in $\mathbb{R}^+ \times \Omega$; $P(V)$ is generated by all path-continuous stochastic processes that are adapted to $\underline{\sigma}(V)$.

R_N will denote the covariance function of (N_t), H_N its RKHS (reproducing Kernel Hilbert space) with inner product $<\cdot,\cdot>_N$, and R_N the covariance operator of (N_t) in $L_2[0,1]$. Range $(R_N^{1/2})$ is a separable Hilbert space, isomorphic to H_N, under the inner product, $(u,g)_N = \sum_n <u,e_n><g,e_n>/\lambda_n$, where $<\cdot,\cdot>$ is the $L_2[0,1]$ inner product, $\{\lambda_n, n \geq 1\}$ are the non-zero eigenvalues of R_N, and $\{e_n, n \geq 1\}$ are associated o.n. eigenvectors.

$\mathbb{R}^{[0,1]}$ is the space of real-valued functions on $[0,1]$; $R^{[0,1]}$ is the Borel σ-field generated by the cylinder sets $\{f$ in $R^{[0,1]}$: $(f(t_1),...,f(t_n)) \in A^n\}$, $n < \infty$, A^n a Borel set in $\mathbb{R}^n$.

For a scalar stochastic process (V_t), υ_V is the probability induced on $\mathbb{R}^{[0,1]}$ by (V_t). If (V_t) has paths belonging a.s. to $L_2[0,1]$, then μ_V will denote the probability induced by the path map on the Borel σ-field of $L_2[0,1]$. If υ_1 and υ_2 are two probabilities on the same σ-field, then $\upsilon_1 << \upsilon_2$ means that υ_1 is absolutely continuous with respect to υ_2.

ABSOLUTE CONTINUITY

Theorem 1 [8]: Let (V_t) be a stochastic process independent of (N_t). Suppose that (Y_t) is

a process such that $\upsilon_Y << \upsilon_N$.

If (Y_t) is adapted to $\underline{\sigma}(N)$ v $\underline{\sigma}(V)$, then $Y_t = S_t + N_t^*$ a.e. dP for each fixed t in $[0,1]$, where (N_t^*) has the same finite dimensional distributions as (N_t), and is adapted to $\underline{\sigma}(Y)$. $N_t^* = \sum_{i=1}^{M} \int_0^t F_i(t,s) dB_i^*(s)$ a.e. dP, each fixed t in $[0,1]$, where the B_i^*s are mutually independent zero-mean Gaussian processes, (B_t^*) has the same law as (B_t), and $\underline{\sigma}(B^*) = \underline{\sigma}(N^*)$. Moreover,

$$S_t = \sum_{i=1}^{M} \int_0^t F_i(t,s) \phi_i(s) d\beta_i(s), \tag{2}$$

where $(\phi_i(t)), i \leq M$, is a stochastic process that is $\underline{\sigma}(Y)$-predictable and has paths a.s. in $L_2[\beta_i]$.

If both (N_t) and (Y_t) have continuous paths, then Theorem 1 can be strengthened. In that case, let P_N' and P_Y' be the induced measures on C. Then $\upsilon_Y << \upsilon_N <=> P_Y' << P_N' <=> \mu_Y << \mu_N$.

<u>Theorem 2</u> [8]: Let (V_t) be a stochastic process independent of (N_t). Suppose that (S_t) is a stochastic process adapted to $\underline{\sigma}(N)$ v $\underline{\sigma}(V)$ and with paths a.s. in H_N.

(1) If $X_t = S_t + N_t$ a.e. dP, for each fixed t in $[0,1]$, then $\upsilon_X << \upsilon_N$.

(2) If $X_t = S_t + N_t$ a.e. dtdP, then $\mu_X << \mu_N$.

<u>Likelihood Ratio</u>

Suppose that (Y_t) satisfies the measurability assumption in Theorem 1, and that $\upsilon_Y << \upsilon_N$. Define a vector process (Z_t) with paths a.s. in C_0^M by

$$Z_i(t) = \int_0^t F_i(t,s) \phi_i(s) d\beta i(s) + B_i(t) \tag{3}$$

where ϕ_i is defined in Theorem 1. In this case, $P_Z << P_B$ [14].

Theorem 3 [8]: Suppose that (Y_t) satisfies the sufficient conditions of Theorem 2. Then

$$\frac{d\,\upsilon_Y}{d\,\upsilon_N}(x) = \int_{C_0^M} [dP_Z/dP_B](y)\, dP_B|_{N=x}(y)$$

a.e. $d\,\upsilon_N(x)$, where $P_B|_{N=x}$ is the conditional measure of B given $N = x$. If (S_t) is defined as in Theorem 1, and $Y = X + N$, then

$$\frac{d\,\mu_Y}{d\,\mu_N}(x) = \int_{C_0^M} [dP_Z/dP_B](y)\hat{P}(x, dy)$$

a.e. $d\,\mu_N(x)$, where $\hat{P}$ is a transition probability on $L_2[0,1] \times C^M$, and $\hat{P}(x,\cdot) \perp P_B$ a.e. $d\,\mu_N(x)$. Moreover, $\hat{P}(x,\cdot)$ is a point mass on C_0^M, giving probability one to $\{m(y)\}$, where

$$[m_i(y)](t) = \sum_n \langle y, e_n\rangle \langle f_t^i, e_n\rangle / \lambda_n \text{ with}$$

$$f_t^i(s) = \int_0^t F_i(s,u)\,d\beta_i(u).$$

From Theorem 3, one can obtain $d\,\upsilon_Y/d\,\upsilon_N$ and $d\,\mu_Y/d\,\mu_N$ from dP_Z/dP_B. Since (B_t) is a vector process with components that are mutually independent continuous-path Gaussian martingales w.r.t. $\underline{\sigma}(N)$, dP_Z/dP_B can be obtained from well-known results [14].

END
DTIC
9-86